Attorney's Docket No. 96790P517

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application for:	
Eiji Murakami et al.	Examiner: Shyue Jiunn Hwa
Serial No. 10/563,311	Art Unit: 2163
Filed: 12-30-2005	Confirmation No.: 6897
For: SENTENCE CLASSIFICATION DEVICE AND METHOD	

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

PROPOSED AMENDMENT AND RESPONSE TO OFFICE ACTION

In response to the Office Action mailed January 22, 2009 in connection with the above referenced patent application, the Applicants respectfully request reconsideration in view of the following amendments and remarks.

Amendments to the claims begin on page 2.

Remarks begin on page 6.

CLAIMS

1. (Currently Amended) A sentence classification device comprising:

a processor;

a memory coupled to the processor to store a term list having a plurality of terms each comprising not less than one word;

a DT matrix generation module for generating a DT matrix two-dimensionally expressing a relationship between each document contained in a document set and said each term;

a DT matrix transformation module for generating a transformed DT matrix having respective clusters, each cluster having one or more blocks of associated documents, by transforming the DT matrix obtained by said DT matrix generation module on the basis of a DM decomposition method used in a graph theory to enable document classification without having to pre-select cluster categories;

a classification generation module for generating classifications associated with the document set on the basis of a relationship between each cluster on the transformed DT matrix obtained by said DT matrix transformation module and said each document classified according to the clusters, wherein the classification generation module comprises a virtual representative document generation module for generating a virtual representative document, for each cluster on a transformed DT matrix, from a term of each document belonging to the cluster;

a large classification generation module for generating a large classification of documents from each document in a bottom-up manner by repeatedly performing, at each DT matrix transformation, said DM decomposition method used to hierarchically cluster documents by setting said DT matrix generated by said DT matrix generation module in an initial state, causing said virtual representative document generation module to generate a virtual representative document for each cluster on a transformed DT matrix generated from the DT matrix by said DT matrix transformation module, generating a new DT matrix used for next hierarchical clustering processing by adding the virtual representative document to the transformed DT matrix and deleting documents belonging to the cluster of the virtual representative document from the transformed DT matrix, and outputting, for said each cluster, information associated with the documents constituting the respective cluster as large classification data of one or more cluster categories;

a term list edition module for adding or deleting an arbitrary term with respect to the term list;

Deleted: and

an index generation module for making said DT matrix generation module generate DT matrices by using term lists before and after edition by said term list edition module, and generating and outputting an index indicating validity of the edition from the DT matrices,

a large classification label generation module for, if a virtual representative document is contained in a given cluster of the respective clusters obtained by the clustering processing, generating a label of the given cluster on which the virtual representative document is based from a term strongly connected to the virtual representative document subsequent to classification of the documents into the respective clusters,

wherein said large classification generation module terminates repetition of the clustering processing when no cluster is obtained from the transformed DT matrix in the clustering processing.

2. (Cancelled)

3. (Currently Amended) A sentence classification device according to claim 1, characterized by further comprising label generation module for outputting each term strongly connected to each document belonging to said arbitrary cluster as a label indicating a classification of the cluster.

4. (Currently Amended) A sentence classification device according to claim 1, further comprising a document organization module for sequentially outputting documents belonging to said arbitrary cluster or all documents in an arrangement order of the documents in the transformed DT matrix.

Deleted: characterized by

5. (Currently Amended) A sentence classification device according to claim 1, further comprising a summary generation module for outputting, as a summary of said arbitrary document, a sentence of sentences constituting the document which contains a term strongly connected to the document.

Deleted: characterized by

6-9. (Cancelled)

10. (Currently Amended) A sentence classification method comprising:

a DT matrix generation step of generating a DT matrix two-dimensionally using a computer, the DT matrix expressing a relationship between each document contained in a document set and each term of a term list having a plurality of terms each comprising not less than one word;

a DT matrix transformation step of generating a transformed DT matrix having respective clusters, each cluster having one or more blocks of associated documents, by transforming the DT matrix on the basis of a DM decomposition method used in a graph theory to enable document classification without having to preselect cluster categories;

a classification generation step of generating classifications associated with the document set on the basis of a relationship between each respective cluster on the transformed DT matrix and said each document classified according to the respective clusters, wherein the classification generation step comprises the virtual representative document generation step of generating a virtual representative document, for each respective cluster on a transformed DT matrix, from a term of each document belonging to the respective cluster;

a large classification generation step of generating a large classification of documents from each document in a bottom-up manner by repeatedly performing, at each DT matrix transformation, said DM decomposition method used to hierarchically cluster documents by setting said DT matrix generated in said DT matrix generation step in an initial state, generating a virtual representative document in the virtual representative document generation step for each respective cluster on a transformed DT matrix generated from the DT matrix in the DT matrix transformation step, the step of generating a new DT matrix used for next hierarchical clustering processing by adding the virtual representative document to the transformed DT matrix and deleting documents belonging to the cluster of the virtual representative document from the transformed DT matrix, and the step of outputting, for said each respective cluster, information associated with the documents constituting the respective cluster as large classification data of one or more cluster categories; and

adding or deleting an arbitrary term with respect to the term list; and the step of generating DT matrices by using term lists before and after edition, and generating and outputting an index indicating validity of the edition from the DT matrices, and

large classification label generation step of, if a virtual representative document is contained in a given cluster of the respective clusters obtained by the clustering processing, generating a label of the given cluster on which the virtual representative document is based from a term strongly connected to the virtual representative document subsequent to classification of the documents into the respective clusters,

wherein in the large classification generation step, repetition of the clustering processing is terminated when no cluster is obtained from the transformed DT matrix in the clustering processing.

11. (Cancelled)

12. (Currently Amended) A sentence classification method according to claim 10, characterized by further comprising the step of outputting each term strongly connected to each document belonging to said arbitrary cluster as a label indicating a classification of the cluster.

13. (Currently Amended) A sentence classification method according to claim 10, further comprising the step of sequentially outputting documents belonging to said arbitrary cluster or all documents in an arrangement order of the documents in the transformed DT matrix.

Deleted: characterized by

14. (Currently Amended) A sentence classification method according to claim 10, further comprising the step of outputting, as a summary of a document, a sentence of sentences constituting said arbitrary document which contains a term strongly connected to the document.

Deleted: characterized by

15-18. (Cancelled)

REMARKS

Claims 1, 3-5, 8-10, 12-14, 17 and 18 were presented and examined. In response to the Office Action, Claims 1, 3-5, 10 and 12-14 are amended, Claims 8-9 and 17-18 are cancelled and no claims are added. Claims 2, 6-7, 11, and 15-16 were cancelled previously. The Applicants respectfully request reconsideration in view of the following remarks and amendments.

I. Claims Rejected Under 35 U.S.C. § 101

Claims 1, 3-5 and 8-9 are again rejected under 35 U.S.C. § 101 as raising a question of "whether the claim is directed merely to an abstract idea that is not tied to a technological art, environment or machine … producing a concrete, useful, and tangible result." (Office Action dated January 22, 2009, page 3.)

The criteria for determining statutory subject matter under §101 have been replaced by the machine-or-transformation test as the "definitive test to determine whether a process claim is drafted narrowly enough to encompass only the particular application of a fundamental principle rather than to pre-empt the principle." *See* In re Bilski, N0 2007-1130, maj. Slip op. (Fed. Cir. 2008). Under *In re Bilski*, a process claim is patent-eligible subject matter if: (1) it is tied to a particular machine or apparatus, or (2) it transforms a particular article into a different state or thing. (*See* supra.)

We submit that the Examiner is misapplying the test for statutory subject matter to non-process claims such as the sentence classification device of Claim 1. Furthermore, Claims 1 and 11 recite the transformation of a term list and a document set into large classification data of documents into non-preselected clusters (categories) as well as label assignment to indicate cluster classification subsequent to classification of documents into the respective clusters. As a result, Claims 1 and 11 provide transformation of documents into a classification hierarchy of cluster categories. As a result, please reconsider and withdraw the §101 rejection of Claims 1, 3-5 and 8-9.

Regarding Claims 10 and 12-14, Claim 10 is tied to a sentence classification device recites a computer to perform the DT matrix generation step. As a result, Claims 10 and 12-14 are directed to statutory subject matter. *In re Bilski*. Therefore, please reconsider and withdraw

the §101 rejection of Claims 10, 13, 14, 17 and 18. In addition, for at least the reasons provided above, please reconsider and withdraw the § 101 rejection of Claims 1-10 and 16-28, which recite analogous claim features to Claim 11.

II. Claims Rejected Under 35 U.S.C. § 103

Claims 1, 3-5, 8-10, 12-14 and 17-18 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over U.S. Patent 7,024,400 issued to Tokuda et al. ("Tokuda") in view of U.S. Patent 6,067,259 issued to Handa, et al. ("Handa"), U.S. Publication No. 2001/0037324 to Agrawal ("Agrawal"), and U.S. Publication 2004/0205457 to Bent, et al. ("Bent"). Applicants respectfully traverse this rejection.

Claim 1 recites:

1. A sentence classification device comprising:

...

a DT matrix transformation module for generating a transformed DT matrix having respective clusters, each cluster having one or more blocks of associated documents, by transforming the DT matrix obtained by said DT matrix generation module on the basis of a DM decomposition method used in a graph theory to enable document classification without having to preselect cluster categories;

...

large classification generation module for generating a large classification of documents from each document in a bottom-up manner by repeatedly performing, at each DT matrix transformation, said DM decomposition method used to hierarchically cluster documents by setting said DT matrix generated by said DT matrix generation module in an initial state, causing said virtual representative document generation module to generate a virtual representative document for each respective cluster on a transformed DT matrix generated from the DT matrix by said DT matrix transformation module, generating a new DT matrix used for next hierarchical clustering processing by adding the virtual representative document to the transformed DT matrix and deleting documents belonging to the cluster of the virtual representative document from the transformed DT matrix,

a large classification label generation module for, if a virtual representative document is contained in a given cluster of the respective clusters obtained by the clustering processing, generating a label of the given cluster on which the virtual representative document is based from a term strongly connected to the virtual representative document subsequent to classification of the documents into the respective clusters,

wherein said large classification generation module terminates repetition of the clustering processing when no cluster is obtained from the transformed DT matrix in the clustering processing. (Emphasis added.)

While Applicant's argument here is directed to the cited combination of references, it is necessary to first consider their individual teachings, in order to ascertain what combination (if any) could be made from them.

Claim 1 features repeated execution of a graph theory called DM decomposition to hierarchically cluster documents. It does not rely on the following known techniques or procedures to executed document clustering: use of statistical Fisher values; use of structure (e.g. hyperlink characters); use of Latent Semantic Indexing (LSI) as a mathematical processor; use of DLSI; or expressing a plurality of documents as vectors in an n-dimensional vector space.

Tokuda prepares, as a supervisor, a document contained in a preselected cluster to determine a center of the cluster in advance. Then, for an unclassified document: 1) a difference between the unclassified document and a center of each cluster to which the document belongs is expressed in a matrix to calculate an intra-DLSI from the matrix; and 2) a difference between the unclassified document and a center of each cluster to which the document does not belong is expressed in a matrix to calculate an extra-DLSI. From the resultant two DLSIs, it is determined to which cluster each document belongs.

In contrast, Claim 1 is directed to an unsupervised classification technique for clustering ordinary unclassified texts into a hierarchical structure without using a supervisor. Meanwhile, Tokuda is directed to a supervised classification technique in which documents belonging to clusters to be used as a number of classification targets are prepared first and then the classification is performed by determining to which of the thus prepared clusters an unclassified document belongs, thus representing a technique that is obviously different from the unsupervised classification technique according to Claim 1.

Furthermore, Claim 1 describes a basic concept of repeatedly executing a graph theory called "DM decomposition" for clustering documents. Conversely, Tokuda executes a document classification process based on a mathematical process called Latent Semantic Indexing (LSI).

Handa is directed to a method and device for repairing memory faults in the course of manufacturing. Handa includes memory cells as an essential component element thereof and makes use of a DM decomposition method, which is a method of calculation used in a graph theory, which is a branch of mathematics. Handa, however, does not teach either repeatedly executing the DM decomposition method or performing document classification hierarchically by the repeated execution of the DM decomposition method.

Claim 1 is distinct from Handa not only in the target data (DT matrix or memory cells) to which the DM decomposition method is to be applied, but also in the objective of the invention (clustering of documents or repair of memory faults). Further, the structure of the target data processed in Claim 1 differs from the structure of data to be processed in Handa. Yet further, Claim 1 employs a data processing procedure that is quite different from the procedure (algorithm) for processing data associated with Handa (with reference to FIG. 1 of Handa).

As described in Claim 1, a DM decomposition method (graph theory) is repeatedly executed to perform the document classification (clustering) to enable document classification without having to preselect clusters. Referring to Bent, Bent describes a plurality of documents expressed in vectors in an n-dimensional vector space (matrix) based on their characteristic quantities (e.g., usage of words in each document) so as to cluster the documents based on proximity in the vector space. A summary of each document comprising a plurality of sentences is generated based on information on words commonly had by the sentences belonging to a cluster.

In contrast, Claim 1 describes a basic concept of repeatedly executing a graph theory called "DM decomposition" for clustering documents to enable document classification without having to preselect cluster categories. Conversely, Bent executes document clustering in the course of generating a summary of each document. The method of Bent therefore comprises expressing each of a plurality of documents in vectors in an n-dimensional vector space based on a characteristic quantity of each document to thereby cluster the documents based on proximity in the vector space. No hierarchical clustering is performed in Bent. Bent does not teach the DM decomposition method itself, and Handa does not teach repeatedly executing the DM

decomposition method. Therefore, the combination of Bent and Handa does not teach repeatedly executing a DM decomposition method to perform document classification, as in Claim 1.

Claim 1 further recites outputting each term strongly connected to each document belonging to the respective cluster as a label indicating a classification of the cluster into a cluster category, subsequent to classification of the documents into the respective clusters. In Tokuda, the cluster is given or preselected and a DLSI classifier is used to determine to which cluster an unclassified sentence belongs. Handa is directed to a technique for repairing memory faults by using a DM decomposition method. In Agrawal, a text which has been preorganized into a hierarchical structure is used as a supervisor to determine to which class an unclassified text belongs. Bent teaches to generate a summary of each document comprising a plurality of sentences based on information on words commonly had by the sentences belonging to a cluster.

Although the Examiner has observed that the presented invention would have been obvious based on a combination of the cited references, the cluster (or a category) is given or preselected in all of Tokuda, Handa, and Agrawal. In addition, for determining to which cluster (category) an unclassified document belongs, classification is performed based on a mathematical processing or supervisor information. Even if the above three references are further combined with Bent, nothing could be achieved other than an idea to generate a summary of each document comprising a plurality of sentences based on information on words commonly had by the sentence belonging to a classified cluster.

In contrast, Claim 1 features classifying documents in a stepwise manner by repeatedly executing a graph theory called "DM decomposition." As described by Claim 1, documents are hierarchically classified in a stepwise manner by repeatedly executing the DM decomposition to transform a graph matrix. Thus, the cluster (or category) is not given in advance (not preselected) according to Claim 1. Neither is supervisor information generated to classify an unclassified sentence. Further, Claim 1 hierarchically classifies documents by repeatedly performing the DM decomposition, without determining to which preselected hierarchy a document belongs.

In light of the foregoing, it cannot be concluded that the constitutional arrangement of Claim 1 could have been realized, nor obvious, based on any combination of Tokuda, Glover, Agrawal, and/or Bent.

Hence, no combination of Tokuda, Glover, Agrawal, and Bent can teach a large classification generation module that repeatedly performs, at each DT matrix transformation, a DM decomposition method used in a graph theory to hierarchically cluster documents to enable document classification without having to preselect cluster categories, as in Claim 1.

Furthermore, the Examiner has not identified, and we are unable to discern any portion of Tokuda, Glover, Agrawal, and Bent that teaches or suggest a label generation module for outputting each term strongly connected to each document belonging to the respective cluster as a label indicating a classification of the cluster into a cluster category, subsequent to classification of the documents into the respective clusters; and a display to show at least the labels, the indexes, and the large classification data of the cluster categories.

For each of the above reasons, Claim 1 and all claims which depend from Claim 1 are patentable over the cited references. Therefore, please reconsider and withdraw the § 103(a) rejection of Claims 1, 3-5, 8-10, 12-14, and 17-18.

Each of the Applicants other independent claims include limitations similar to those discussed above. Therefore, all of the Applicants other independent claims, and all claims which depend on them, are patentable over the cited art for similar reasons. Consequently, Applicants respectfully request that the Examiner reconsider and withdraw the §103 rejection of Claims 10, 12-15 and 17-18.

DEPENDENT CLAIMS

In view of the above remarks, a specific discussion of the dependent claims is considered to be unnecessary. Therefore, Applicant's silence regarding any dependent claim is not to be interpreted as agreement with, or acquiescence to, the rejection of such claim or as waiving any argument regarding that claim.

CONCLUSION

In view of the foregoing, it is submitted that the pending claims patentably define the subject invention over the cited references of record, and are in condition for allowance and such action is earnestly solicited at the earliest possible date. If the Examiner believes a telephone conference would be useful in moving the case forward, he is encouraged to contact the undersigned at (310) 207-3800.

If necessary, the Commissioner is hereby authorized in this, concurrent and future replies, to charge payment or credit any overpayment to Deposit Account No. 02-2666 for any additional fees required under 37 C.F.R. §§1.16 or 1.17, particularly, extension of time fees.

Respectfully submitted,

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP

Dated: May 2, 2009

By: _____
Joseph Lutz, Reg. No. 43,765

1279 Oakmead Parkway
Sunnyvale, CA 94085-4040
Telephone (310) 207-3800
Facsimile (408) 720-8383

CERTIFICATE OF TRANSMISSION
I hereby certify that this correspondence is being submitted electronically via EFS Web to the United States Patent and Trademark Office on May 2, 2009.

_____
Si Vuong